jh210026-NAH

Innovative Multigrid Methods II

Akihiro Fujii (Kogakuin University)

Abstract

Multigrid method is known as one of the most efficient and scalable methods using the supercomputers. It is applicable to partial differential equations with fine and coarse grids, linear matrix problems, and parallel time integration. Our project studies multigrid related themes in 3-year plan. This report is for the second year. Although we continue studying the area, we've conducted researches in 2021, including the topics of precision optimization with SELL-C-σ sparse matrix format, computation and communicaiton overlapping in sparse matrix vector product, weak scaling evaluation of AMG solver, new parareal method for explicit time-marching applications, and scalability enhancement method for Parallel in Space Time solvers.

1 Basic Information

1.1 Collaborating JHPCN Centers

Tokyo, Nagoya

1.2 Research Areas

- Very large-scale numerical computation

1.3 Roles of Project Members

Akihiro Fujii[3] :AMG, PinST
Kengo Nakajima[2,4] : (Co-PI) Application, GMG, PinST, AMG
Matthias Bolten[8] : (Co-PI) GMG, PinST, AMG
Masatoshi Kawai[2] : GMG, AMG
Akihiro Ida[13] : GMG, AMG
Gerhard Wellein[9] : GMG, AMG
Christie Alappat[9] : GMG, AMG
Martin Schreiber[11] : GMG, AMG
Tetsuya Hoshino[2] : GMG, AMG
Satoshi Ohshima[6] : GMG, AMG
Toshihiro Hanawa[2] : GMG, AMG
Osni Marques[10] : GMG, AMG
Kenji Ono[5] : PinST
Mikio Iizuka[5] : PinST
Takeshi Iwashita[1] : AMG, PinST
Yasuhito Takahashi[7] : PinST
Robert Speck[12] : PinST
Atsuhiro Miyagi[14]: PinST
Teruo Tanaka[3] :AMG, PinST
Alexander T.Magro[2] : GMG
Ryo Yoda[2] : PinST, AMG
Yen-Chen Chen[2] : PinST
Gayatri Caklovic[12]:PinST

1: Hokkaido U., 2: U. Tokyo, 3: Kogakuin U., 4: RIKEN R-CCS, 5: Kyushu U., 6: Nagoya U., 7: Doshisha U., 8: U. Wuppertal*, 9: FAU*, 10: LBNL, USA, 11: Technical U. of Munich*, 12: Juelich Supercomputing Centre*, 13: JAMSTEC, 14: Taisei Corp. *:Germany

2 Purpose and significance of Research

Multigrid method is a promising approach for large-scale computing in exa-scale era. We develop robust and efficient parallel multigrid methods for both of geometric multigrid (GMG) and algebraic multigrid(AMG), focusing on robust and efficient smoothers, and hierarchical methods which are proposed and developed by ourselves. We have original PinST methods. Developed methods will be implemented as a numerical library and it will be applied to various types of applications. It is expected to provide outstanding performance for large-scale real-world applications.

3 Significance as JHPCN Joint Research Project

Multigrid method is scalable and used in many fields. It is known as one of the most efficient linear solvers. It can also be applied to parallel time integration problems, which exploits parallelism in time dimension. Our research project has original codes and algorithms as written in Research purpose. Therefore, research papers and codes from the project will enhance the efficiency of the multigrid solver, and will help many researchers exploit parallelism in time direction.

Our research focuses on hierarchical algorithms and their performance on supercomputers. Thus, availability of supercomputers with different kinds of architectures helps us verify the codes we are developing. In addition, a JHPCN joint research project offers collaborative research opportunity with JHPCN members who have expertise knowledge in various application fields. Our project members include international experts in Germany, US, and Japan on multigrid methods and PinST. We are sure that this JHPCN joint research project promotes the international collaborative activity with JHPCN members.

4 Outline of Research Achievements up to FY2021

This project is set as a 3-year project, and this report is for the second year. Main presentation and publication of the second year is listed in section 7. For GMG, we have conducted precision optimization with SELL-C-σ in [GMG_1], and computation and communication overlapping in [GMG_4]. For AMG, we have checked our AMG solver's scaling performance with large sized problems with 10^{12} unknowns in [AMG_1]. As for PinST, we studied parareal method for explicit time marching scheme in [PinST_exp_1]. Performance enhancement technique adjusting the balance of parallelism in space and time directions was proposed in [PinST_scl_1].

5 Details of FY2021 Research Achievements

In the 2nd year of our continuing project, we are doing researches including the following items.

1. GMG: Precision optimization with SELL-C-σ
2. AMG: Weak scaling performance test on Supercomputer Flow and Fugaku
3. PinST: New explicit time-marching application
4. PinST: Scalability enhancement for MGRIT
5. Comm: Fast halo communication directly using RDMA interface

Subsections introduce the researches.

5.1 GMG: Parallel multigrid using SELL-C-σ with FP64/32, and communication computation overlapping technique

A paper on parallel multigrid method using SELL-C-σ with FP64/FP32 has been presented at iWAPT workshop in IPDPS 2021 [GMG_1], and there were many useful comments from audience for future works, such as effects of precision on MPI communications. The developed method in [GMG_1] has been applied to Wisteria/BDEC-01 (Odyssey) with A64FX, and performance has been evaluated using up to 4,096 nodes [GMG_2,3]. Configuration of network topology affects the performance of all-to-all-type communications, if the number of nodes is more than 2,048.

In [GMG_4], we studied communication-computation overlapping. Communication overhead is a critical issue when executing sparse matrix solvers on large-scale massively parallel supercomputers. In the previous work, we introduced communication-computation overlapping with dynamic loop

Fi

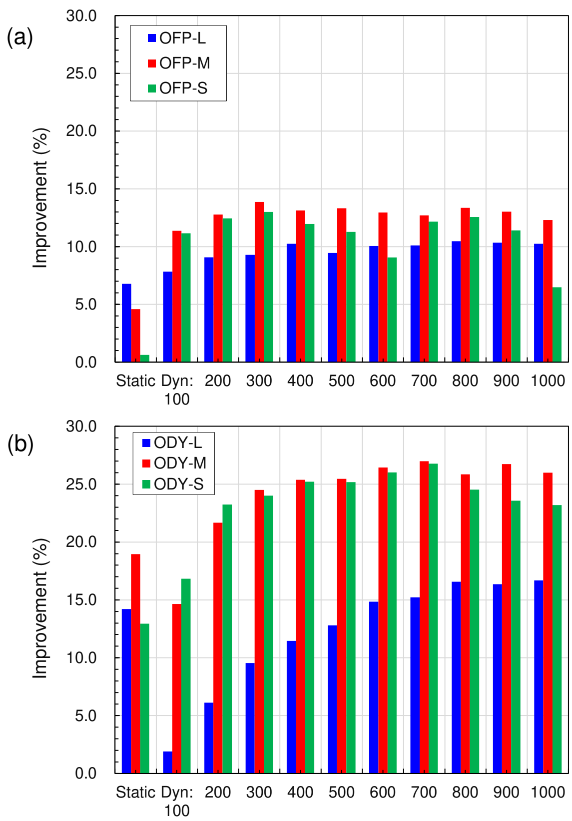

Fig. 1 Improvement by Static and Dynamic (Dyn) CC-Overlapping over "original" implementations at 2,048 nodes (a) OFP with IHK/McKernel, (b) Odyssey

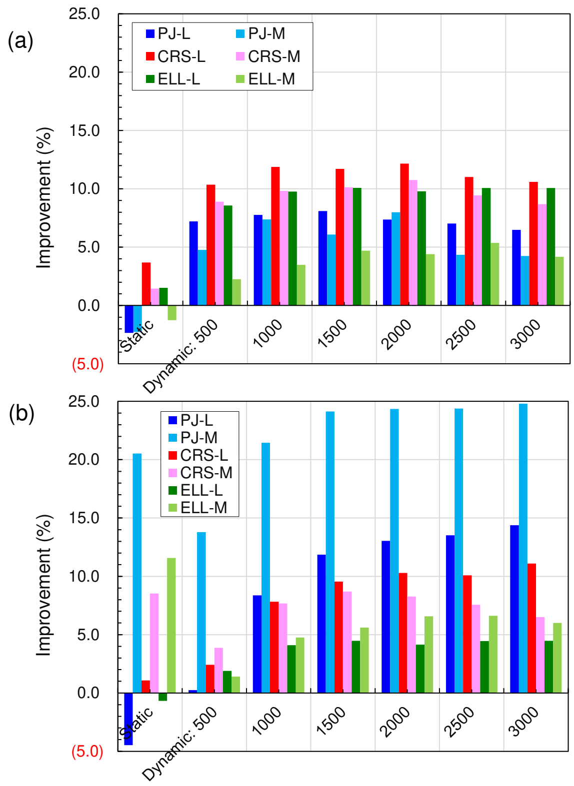

Fig. 2 Improvement by Static and Dynamic CC-Overlapping over "Original" implementation at 1,024 nodes (a) OFP with IHK/McKernel, (b) Odyssey

scheduling of OpenMP to the sparse matrix-vector multiplication (SpMV) process of a parallel iterative solver by Conjugate Gradient (CG) method in a parallel finite element application (GeoFEM/Cube) on multicore and manycore clusters. In the present work, first, we re-evaluated the method on Wisteria/BDEC-01 (Odyssey) (Fujitsu PRIMEHPC FX1000 with A64FX), and a significant performance improvement of 25-30% for parallel iterative solver at 2,048 nodes (98,304 cores) was obtained. Moreover, we proposed a new reordering method for communication-computation overlapping in ICCG solvers for a parallel finite volume application (Poisson3D/Dist), and attained 5-12% improvement at 1,024 nodes of Odyssey.

5.2 AMG: Weak scaling performance test on Supercomputer Flow and Fugaku

We proposed light-weight process aggregation method using parallel C/F splitting algorithm (CLJP). From the numerical tests, it turned out that our CLJP-based method was better than Coarse grid aggregation with ParMETIS library especially in the computing environment with more than 100000 processes. In addition to the new process aggregation method, we improved our implementation for large sized coarse level matrix whose size is over the range of 32-bit integer type.

We evaluated the solver with this CLJP-based coarse grid aggregation on Supercomputer Fugaku using up to about 50000 nodes. Here, we introduce the weak scaling evaluation with Supercomputer Fugaku. Numerical test condition is listed in Table 1. Since more

Table 1 Weak scaling test setting

Domain size of a process	$288 \times 144 \times 144$
Process/thread allocation on 1 node	4 MPI processes 12 Open-MP thrds
CRS matrix size of 1 node	7.8GB 27 ele. per row
Number of nodes	from 768 to 49152 nodes
Max.number of cores and processes	2.35×10^6 cores 1.9×10^5 processes
Max. domain size	1.17×10^{12}DOF

than 50000 nodes are used in the test, 4 processes are assigned to each node. The CPU has 4 CMGs (Core Memory Groups) and 4 MPI-processes per 1 node setting is recommended in general. The problem matrix size was set as about 8GB per node, which occupies a 1/4 each node 32GB memory.

The solver was AMG preconditioned CG method with multi-color Gauss Seidel smoother. Thus it was executed in MPI-OpenMP hybrid model. The convergence criterion was the relative residual 2-norm less than 1.0×10^{-6}.

Fig. 3 and Fig. 4 show total execution time and AMG-CG one iteration time, respectively. The horizontal axis of these figures indicates the problem DOFs in logarithmic scale. The vertical axis indicates the time in seconds. Fig. 3 plots total execution time and multi-level creation (Setup) time. Numbers in the figure correspond to the iteration number for convergence. Total execution time requires only 1.31 times longer for 64 times larger problem. Iteration number also increases from 30 to 37. As a whole, it demonstrates a good weak scaling performance. Fig. 4 plots AMG-CG one iteration time. From smallest to the largest problem sizes, the time difference was only 7 %.

In the largest problem, the number of unknowns at the second coarse level is 4.32×10^{10} which is over the maximum value of 32 bit integer. For such problems, the solver creates distributed matrices with 32-bit integer, and reaches convergence. Hence, it is assumed to be applicable to larger problems, as long as the block row matrix allocated to 1 process is represented in the range of 32-bit integer type.

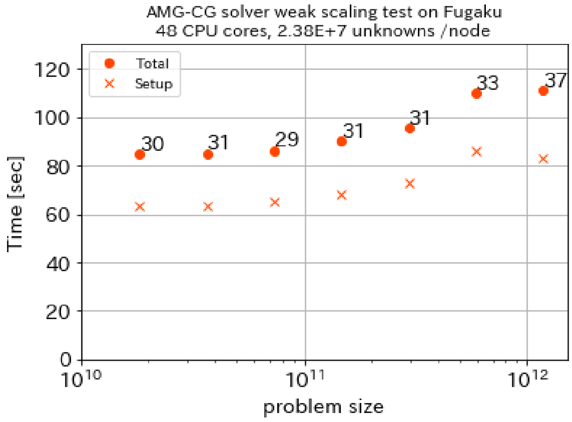

Fig. 3 Weak scaling test on Fugaku

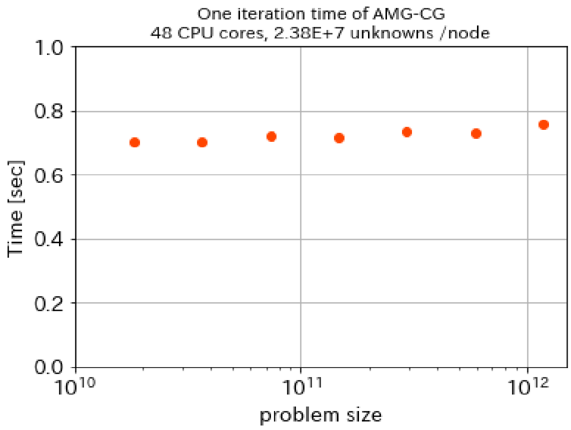

Fig. 4 AMG-CG 1 iteration time

5.3 PinST: explicit time-marching application

Parallel-in-Space/Time (PinST) method for explicit time-marching application has been developed based on parareal method. The method has been applied to 1D convection-diffusion equations, and 2D compressible viscous flow problems with Navier-Stokes equations. Proposed method provides better performance and stability compared to existing methods, such as parareal, and MGRIT. This

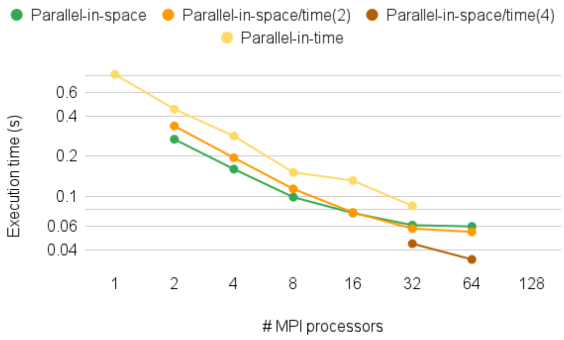

Fig. 5 Execution time comparison of spatial parallelization to parallel-in-space/time for one-dimensional advection example.

is one of the first example where PinST is successful for practical applications with explicit time-marching. The proposed method and results were presented at domestic and international meetings [PinST_exp_2,3], and a paper submitted to ScalA21 workshop in SC21 has been accepted for presentation [PinST_exp_1]. While the method is applied to subsonic flow currently, method for supersonic cases using adaptive mesh refinement is under development. Fig. 5 introduces the performance comparison among Parallel-in-space, Parallel-in-time, and Parallel-in-space/time for one-dimensional advection example.

5.4 PinST: MGRIT scaling enhancement

In the previous work, we apply coarse-grid agglomeration (CGA) to MGRIT to enhance scalability, which gradually shrinks the number of active processes on a coarse grid. While this technique achieved well-scaling performance, it makes more processors idle than usual spatial multigrid methods. Therefore, if these idle processors can be utilized somehow, it will lead to further performance improvement.

These researches [PinST_scl_1,2,4] consider a method to assign these idle processors to the spatial redistributed domain on coarse levels. Then, temporal parallelism decreases by CGA, and spatial parallelism increases with each level. This method aims to accelerate coarse-level spatial solvers.

Numerical experiments were conducted for three-dimensional diffusion problems on a structured grid. We observed the well-scaling performance of up to 512 nodes on OFP and that the proposed method achieves about 1.4 times faster than pure MGRIT. We confirmed the effectiveness of the use of idle processors by CGA.

In addition, the research [PinST_scl_3] investigates the optimized coarse grid operator for time-dependent Stokes and Oseen problems.

5.5 Fast halo Communication

Supercomputers such as Fugaku and FX1000 offer users RDMA communication interface. We investigated fast halo communication using this interface in [Comm_1] and [Comm_2]. One-sided communications such as RDMA write tends to need explicit synchronization in general. We eliminated this explicit synchronization by preparing double buffers, and evaluated the efficiency of the halo communication in typical communication setting.

6 Progress during FY2021 and Future Prospects

Three areas of GMG, AMG and PinST are set in our project. We have made progress in all areas, and the progress is going well. In the next fiscal year, we will organize the research results to publish papers and codes.

7 List of Publications and Presentations

Journal Papers (Refereed)

Proceedings of International Conferences (Refereed)

[PinST_scl_1] Yoda, R., Bolten, M.(+), Nakajima, K., Fujii, A., Assignment of idle processors to spatial redistributed domains on coarse levels in multigrid reduction in time. In HPC Asia ' 22: The International Conference on High Performance Computing in Asia-Pacific Region, Jan 12–14, 2022, Kobe, Japan & Online. ACM, New York, NY, USA (in

press)
[PinST_exp_1] Chen, Y.-C., Nakajima, K., Optimized Cascadic Multigrid Parareal Method for Explicit Time-Marching Scheme, IEEE Proceedings of ScalA21: 12th Workshop on Latest Advances in Scalable Algorithms for Large-Scale Systems, in conjunction with SC21, November 2021 (in press)
[GMG_1] Nakajima, K., Ogita, T., Masatoshi, K., Efficient Parallel Multigrid Methods on Manycore Clusters with Double/Single Precision Computing, IEEE Proceedings of iWAPT 2021 in conjunction with IPDPS 2021, May 2021
[GMG_4] Horikoshi, M., Gerofi, B., Ishikawa, Y., Nakajima, K., Exploring Communication-Computation Overlap in Parallel Iterative Solvers on Manycore CPUs using Asynchronous Progress Control, IXPUG Workshop in conjunction with HPC Asia 2022, 2022
[GMG_5] Kawai, M., Nakajima, K., Low/Adaptive Precision Computation in Preconditioned Iterative Solvers for Ill-Conditioned Problems, The International Conference on High Performance Computing in Asia-Pacific Region (HPC Asia 2022), 2022

Proceedings of International Conferences (Non-refereed)

[GMG_2] 中島 研吾，河合 直聡，Wisteria/BDEC-01(Odyssey) における並列多重格子法ソルバーの開発と性能評価，日本応用数理学会年会 2021，先進的環境における数値計算と関連 HPC 技術，2021 年 9 月
[AMG_1] 藤井 昭宏，田中 輝雄，中島 研吾，SA-AMG 法における軽量な粗格子集約手法と富岳上でのウィークスケーリング性能評価，情報処理学会研究報告（2021-HPC-181）(第 181 回 HPC 研究会)（オンライン，2021 年 9 月 27 日）
[PinST_exp_2] Chen, Y.-C., Nakajima, K., An Efficient Parallel-in-Time Method for Explicit Time-Marching Schemes, PinT 2021: 10th Workshop on Parallel-in-Time Integration, August 2021
[PinST_scl_2] Yoda, R., Bolten, M.(+), Nakajima, K., Fujii, A., Spatial redistribution on the temporal coarse level for Multigrid Reduction in Time, PinT 2021: 10th Workshop on Parallel-in-Time Integration, August 2021
[PinST_scl_3] 依田 凌, 中島研吾, 藤井昭宏, MGRIT の粗格子演算子に対する最適化手法とその線形時間発展 Stokes ・ Oseen 問題への適用, 日本応用数理学会「行列・固有値問題の解法とその応用」研究部会第 32 回研究会. (オンライン, 2021 年 12 月 10 日)
[PinST_scl_4] Yoda, R., Bolten, M.(+), Nakajima, K., Fujii, A., Assignment of Idle Processors to Spatial Redistributed Domains for Multigrid Reduction in Time, SIAM Conference on Parallel Processing for Scientific Computing (PP22), February 23–26, 2022, Virtual Conference, Originally scheduled in Seattle, Washington, U.S.
[GMG_3] 中島 研吾, 河合 直聡, Wisteria/BDEC-01(Odyssey) における大規模前処理付き反復法ソルバーの性能評価, 日本応用数理学会「行列・固有値問題の解法とその応用」研究部会（MEPA），2021 年並列／分散／協調処理に関するサマー・ワークショップ (SWoPP2021), 2021 年 7 月
[PinST_exp_3] Chen, Y.-C.,Nakajima, K., An Effective Parallel-in-Time Method for Explicit Time-Marching Schemes, IPSJ, IPSJ SIG Technical Report（2021-HPC-180-2），2021
[PinST_exp_4] Chen, Y.-C., Nakajima, K., Cascadic Parareal Method for Explicit Time-Marching Schemes, HPC Asia 2022 (3rd Prize, HPC Asia 2022 Best Student Poster Award in Memory of Hiroshi Nakashima), 2022
[Comm_1] Yoshimoto, K., Fujii, A., Tanaka, T., Fast Adjacent Communication with RDMA and Double Buffering, The 20th International Symposium on Advanced Technology (ISAT-20), Lightning Talk session, Nov.23-24, 2021.
[Comm_2] Yoshimoto, K., Fujii, A., Tanaka,

T., RDMA with Double Buffering for Adjacent Communication, The International Conference on High Performance Computing in Asia-Pacific Region (HPC Asia 2022), poster, Jan. 2022.

Published library and relating data

Other (patents, press releases, books and so on)